



Hoang, L. T., Jain, P., Pillay, T. D., Tolosa-Wright, M., Niazi, U., Takwoingi, Y., Halliday, A., Berrocal-Almanza, L. C., Deeks, J. J., Beverley, P., Kon, O. M., & Lalvani, A. (2021). Transcriptomic signatures for diagnosing tuberculosis in clinical practice: a prospective, multicentre cohort study. *Lancet Infectious Diseases*, 21(3), 366-375. [https://doi.org/10.1016/S1473-3099\(20\)30928-2](https://doi.org/10.1016/S1473-3099(20)30928-2)

Publisher's PDF, also known as Version of record

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/S1473-3099\(20\)30928-2](https://doi.org/10.1016/S1473-3099(20)30928-2)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Elsevier at [https://doi.org/10.1016/S1473-3099\(20\)30928-2](https://doi.org/10.1016/S1473-3099(20)30928-2) . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Transcriptomic signatures for diagnosing tuberculosis in clinical practice: a prospective, multicentre cohort study

Long T Hoang, Pooja Jain*, Timesh D Pillay*, Mica Tolosa-Wright, Umar Niazi, Yemisi Takwoingi, Alice Halliday, Luis C Berrocal-Almanza, Jonathan J Deeks, Peter Beverley, Onn Min Kon, Ajit Lalvani



Summary

Background Blood transcriptomic signatures for diagnosis of tuberculosis have shown promise in case-control studies, but none have been prospectively designed or validated in adults presenting with the full clinical spectrum of suspected tuberculosis, including extrapulmonary tuberculosis and common differential diagnoses that clinically resemble tuberculosis. We aimed to evaluate the diagnostic accuracy of transcriptomic signatures in patients presenting with clinically suspected tuberculosis in routine practice.

Methods The Validation of New Technologies for Diagnostic Evaluation of Tuberculosis (VANTDET) study was nested within a prospective, multicentre cohort study in secondary care in England (IDEA 11/H0722/8). Patients (aged ≥ 16 years) suspected of having tuberculosis in the routine clinical inpatient and outpatient setting were recruited at ten National Health Service hospitals in England for IDEA and were included in VANTDET if they provided consent for genomic analysis. Patients had whole blood taken for microarray analysis to measure abundance of transcripts and were followed up for 6–12 months to determine final diagnoses on the basis of predefined diagnostic criteria. The diagnostic accuracy of six signatures derived from the cohort and three previously published transcriptomic signatures with potentially high diagnostic performance were assessed by calculating area under the receiver-operating characteristic curves (AUC-ROCs), sensitivities, and specificities.

Findings Between Nov 25, 2011, and Dec 31, 2013, 1162 participants were enrolled. 628 participants (aged ≥ 16 years) were included in the analysis, of whom 212 (34%) had culture-confirmed tuberculosis, 89 (14%) had highly probable tuberculosis, and 327 (52%) had tuberculosis excluded. The novel signature with highest performance for identifying all active tuberculosis gave an AUC-ROC of 0.87 (95% CI 0.81–0.92), sensitivity of 77% (66–87), and specificity of 84% (74–91). The best-performing published signature gave an AUC-ROC of 0.83 (0.80–0.86), sensitivity of 78% (73–83), and specificity of 76% (70–80). For detecting highly probable tuberculosis, the best novel signature yielded results of 0.86 (0.71–0.95), 77% (56–94%), and 77% (57–95%). None of the relevant cohort-derived or previously published signatures achieved the WHO-defined targets of paired sensitivity and specificity for a non-sputum-based diagnostic test.

Interpretation In a clinically representative cohort in routine practice in a low-incidence setting, transcriptomic signatures did not have adequate accuracy for diagnosis of tuberculosis, including in patients with highly probable tuberculosis where the unmet need is greatest. These findings suggest that transcriptomic signatures have little clinical utility for diagnostic assessment of suspected tuberculosis.

Funding National Institute for Health Research.

Copyright © 2021 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.

Introduction

Tuberculosis is the leading infectious cause of death worldwide.¹ A large proportion of the 10 million new cases reported in 2018 were diagnosed without bacteriological confirmation, indicating the shortcomings of available diagnostic approaches.¹ WHO recommends that new diagnostic tests for tuberculosis be low-cost, easy to use, non-invasive, and achieve high sensitivity and specificity.² The Xpert MTB/RIF assay (Cepheid; Sunnyvale, CA, USA) has transformed diagnosis of pulmonary tuberculosis in resource-poor settings; however, it has poor sensitivity for culture-negative tuberculosis.^{3,4} The low sensitivity of bacteriological

diagnostic tools probably reflects the paucibacillary nature of culture-negative and extrapulmonary tuberculosis, which accounts for a large and growing proportion of tuberculosis cases in low-incidence regions. Moreover, invasive sampling procedures, such as bronchoscopy and biopsy, are expensive and carry risk. Therefore, there is an unmet clinical need for improved rapid diagnostic tools for tuberculosis without invasive sampling, which could improve diagnosis of culture-negative tuberculosis and extrapulmonary tuberculosis, for which improved diagnostic tools are needed most.⁵

Blood-based, host-derived immune response biomarkers reflect an amplified signal to the infecting bacillus, so

Lancet Infect Dis 2021

Published Online
January 25, 2021
[https://doi.org/10.1016/S1473-3099\(20\)30928-2](https://doi.org/10.1016/S1473-3099(20)30928-2)

See Online/Comment
[https://doi.org/10.1016/S1473-3099\(20\)30993-2](https://doi.org/10.1016/S1473-3099(20)30993-2)

*These authors contributed equally

Tuberculosis Research Centre, National Heart and Lung Institute, Imperial College London, London, UK (L T Hoang PhD, P Jain PhD, T D Pillay MBBS, M Tolosa-Wright BSc, L C Berrocal-Almanza PhD, P Beverley DSc, Prof O M Kon MD, Prof A Lalvani DM); Guy's and St Thomas' National Health Service Foundation Trust and King's College London National Institute for Health Research Biomedical Research Centre Translational Bioinformatics Platform, Guy's Hospital, London, UK (U Niazi PhD); Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, Birmingham, UK (Y Takwoingi PhD, Prof J J Deeks PhD); National Institute of Health Research Birmingham Biomedical Research Centre, University Hospitals Birmingham National Health Service Foundation Trust and University of Birmingham, Birmingham, UK (Y Takwoingi PhD, Prof J J Deeks PhD); Bristol Children's Vaccine Centre, University of Bristol, UK (A Halliday PhD); Tuberculosis Service, Imperial College Healthcare National Health Service Trust, London, UK (Prof O M Kon)

Correspondence to: Prof Ajit Lalvani, Tuberculosis Research Centre, National Heart and Lung Institute, Imperial College London, London SW3 6LY, UK
a.lalvani@imperial.ac.uk

Research in context

Evidence before this study

Diagnosis of tuberculosis, particularly extrapulmonary and culture-negative tuberculosis, which account for a large and growing proportion of tuberculosis in low-incidence regions, remains a major unmet clinical need. Blood-based host-response biomarkers, particularly genome-wide transcriptomic signatures, hold promise for improving diagnosis. However, existing signatures were derived in case-control studies comprised exclusively of culture-confirmed pulmonary tuberculosis cases and preselected controls. Their performance in clinical practice in patients with a full range of tuberculosis, including extrapulmonary and culture-negative tuberculosis, is unknown. We searched PubMed for articles published between inception and May 31, 2020, using the search terms “transcriptomic”, “transcriptome”, “tuberculosis”, “TB”. This search yielded articles describing the derivation and validation of 28 relevant diagnostic transcriptomic signatures. All adult studies were case-control design with culture-confirmed pulmonary tuberculosis cases and preselected healthy or other disease controls.

Added value of this study

To our knowledge, this is the first study to derive novel transcriptomic signatures for the diagnosis of tuberculosis in routine clinical practice in patients with the full spectrum of tuberculosis recruited at the point of clinical suspicion. It is also the largest cohort for assessment of this technology for

tuberculosis diagnosis, making the findings reliable as well as generalisable to other low-incidence settings. Among 628 participants, the best-performing cohort-derived novel signature for detecting all active tuberculosis gave an area under the receiver-operating characteristic curve of 0.87 (95% CI 0.81–0.92), sensitivity of 77% (66–87), and specificity of 84% (74–91); previously published signatures were no better at detecting all active tuberculosis. Diagnostic performance in highly probable culture-negative tuberculosis, the subgroup in whom the unmet clinical need is greatest and transcriptomic signatures have not previously been assessed, was no better. None of the signatures achieved optimum or minimum WHO criteria for a diagnostic test for tuberculosis.

Implications of all the available evidence

Transcriptomic signatures have shown promise for tuberculosis diagnosis in numerous case-control studies and meta-analyses. However, their performance when applied at the point of diagnostic assessment in patients presenting in routine practice with the full clinical spectrum of suspected tuberculosis is lower than anticipated. The generalisable evidence presented here suggests that transcriptomic signatures are of little value in the diagnosis of suspected tuberculosis in clinical practice in low-incidence settings. These findings exemplify the need for evaluation of new diagnostic technologies in the clinically relevant target population at the point of initial diagnostic suspicion in routine practice.

might improve diagnostic sensitivity compared with microbiologically based methods.² Whole-blood transcriptomic signatures that provide a broad view of the host response to tuberculosis have shown considerable promise to this end—several studies in low-incidence and high-incidence settings have shown accurate differentiation of tuberculosis from healthy controls, tuberculosis from other diseases, and tuberculosis from latent tuberculosis infection.^{6–12}

These promising studies have two major drawbacks. First, all but one (in young children) were case-control studies comprising patients with confirmed tuberculosis and preselected healthy or non-tubercular disease controls. Such studies do not reflect routine clinical practice. Typically, suspicion of tuberculosis needs confirmation or exclusion in patients with a wide range of presenting conditions, which might mimic tuberculosis, and who often have concomitant latent tuberculosis infection.^{13–15} Second, these studies were predominantly restricted to patients with culture-confirmed pulmonary tuberculosis, and patients with extrapulmonary tuberculosis and culture-negative tuberculosis were excluded. However, these subgroups have the greatest unmet need for an improved diagnostic test. Consequently, there is a pressing need to derive and validate novel transcriptomic signatures in routine clinical practice cohorts representing the full

clinical spectrum of tuberculosis and non-tuberculosis differential diagnoses. Such evidence is essential for clinicians and health-care systems to determine whether to deploy this genomic approach as a diagnostic service, and if so, how. We aimed to address this evidence gap in a large multicentre prospective cohort recruited in routine practice designed specifically to assess novel diagnostic tests for tuberculosis in low-incidence settings.

Methods

Study design and participants

We embedded an analysis of transcriptomic signatures within a prospective cohort study in routine clinical practice in England (the IGRAs in Diagnostic Evaluation of Active TB [IDEA] study; approved by Camden and Islington National Research Ethics Committee, reference 11/H0722/8), which was designed to assess the clinical utility of existing and second-generation interferon- γ release assays (IGRAs) in the diagnostic evaluation of tuberculosis, as previously reported.¹⁶ The present study (Validation of New Technologies for Diagnostic Evaluation of Tuberculosis [VANTDET]), funded by the National Institute for Health Research Efficacy, Mechanisms and Evaluation programme (EME 12/65/27), builds on the unique value of the IDEA cohort to evaluate the real-life performance of new technologies, including

transcriptomic signatures, for the diagnosis tuberculosis on samples deposited in the IDEA biobank (respiratory infections tissue bank 07/H0712/85+5) as patients were enrolled.

In the IDEA study, adults initially suspected of having tuberculosis by local attending clinicians in inpatient and outpatient infectious diseases or respiratory secondary care services were recruited by research nurses at the point of diagnostic investigation at ten National Health Service Trust hospitals in England (appendix p 16).¹⁶ Patients younger than 16 years and those unable to give informed consent were not recruited. After recruitment, participants were excluded from this study if they had not provided consent to genomic analysis.

The study was approved by the Camden and Islington National Research Ethics Committee. Participants provided written informed consent.

Procedures

After initial enrolment, blood for IGRAs and transcriptomic signature derivation was sampled and data collected in case report forms. Blood samples were taken for transcriptomic signature analysis before treatment initiation, and the presence or absence of tuberculosis was verified according to stringent, predefined previously validated criteria.^{16,17} 35 mL of blood was collected. RNA was extracted from PAXGene tubes using a blood RNA kit (PreAnalytiXL; Hombrechtikon, Switzerland). RNA quality control was done using both the NanoDrop 8000 (ThermoFisher; Waltham, MA, USA) and the Agilent 2100 Bioanalyzer (Agilent; Santa Clara, CA, USA) instruments. Global transcriptomic profiles of the participants were characterised using whole-genome-wide microarray Illumina (San Diego, CA, USA) gene expression platforms (HT12 version 4 bead chip). Microarray data were submitted to the Gene Expression Omnibus repository with accession number GSE144127. The detailed protocol has been published previously.¹⁸ Scientists who did RNA extraction and microarray analysis were masked to patients' final diagnoses by blocking their access to the clinical database. Participants were excluded at this stage if their blood sample contained low quality or quantity of RNA, had a missing sample date, their RNA was not detected, or showed outlier values on microarray. Outlier values were defined as falling outside the 95% confidence ellipse on principle component analysis.

Participants were followed at 2 months, 6 months, and 12 months, with data collected on investigations, test results, diagnoses, and response to tuberculosis treatment. Participants with a definitive non-tuberculosis diagnosis were not followed further, because they had reached a diagnostic endpoint. Participants lost to follow-up or who withdrew before 6 months and before a definitive non-tuberculosis diagnosis was reached were excluded.

Patients were investigated and treated by respiratory or infectious disease physicians in routine clinical practice. After the 12-month study follow-up, with all clinical,

radiological, histological, and microbiological data available, participants received a definitive diagnostic categorisation. These outcomes were determined by consensus across a panel of five expert respiratory and infectious disease clinicians acting as part of the IDEA study team who were masked to the results of transcriptomic analyses and IGRAs, as previously described.¹⁶ The panel used a predefined validated diagnostic classification,^{16,17} including clinical, radiological, and microbiological data. Category 1, culture-confirmed tuberculosis, was defined as the combination of suggestive clinical and radiological findings and positive microbiological culture for *Mycobacterium tuberculosis*. Category 2, highly probable tuberculosis, was defined as the combination of all of the following stringent criteria: clinical and radiological features highly

See Online for appendix

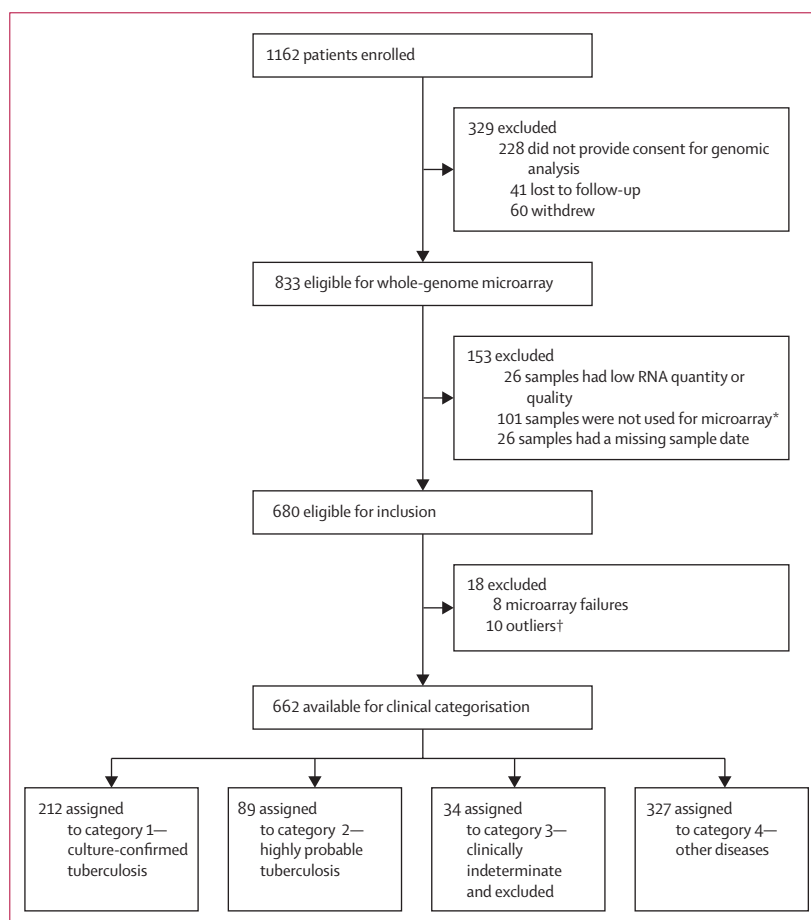


Figure 1: Study profile

*Recruitment was extended from Aug 31, 2013 (as used and reported in Whitworth et al¹⁶), to Dec 31, 2014, to maximise HIV-positive tuberculosis cases to have sufficient numbers to compare the technologies being evaluated in HIV-positive individuals. However, by Dec 31, 2014, this strategy only delivered 20 cases of HIV-positive tuberculosis, still fewer than the number required to meaningfully compare the subgroups. Therefore, 25 HIV-positive participants with tuberculosis who were excluded were randomly selected from the total number of 126 such cases to form a balanced group relative to the 20 HIV-positive tuberculosis cases, meaning that 101 HIV-positive participants with tuberculosis were excluded at this stage. This study included eight HIV-positive patients recruited during this extension period not included in Whitworth et al,¹⁶ one of whom had tuberculosis.

†Defined as falling outside the 95% confidence ellipse on principle component analysis.

	Culture-confirmed tuberculosis (n=212)	Highly probable tuberculosis (n=89)	Clinically indeterminate (n=34)	Other diseases (n=327)
Age, years				
Median (IQR)	32 (26–42)	36 (28–45)	45 (27–57)	43 (33–56)
Range	16–81	18–76	16–79	17–87
Sex				
Male	149 (70%)	45 (51%)	13 (38%)	189 (58%)
Female	62 (30%)	44 (49%)	21 (62%)	138 (42%)
Ethnic origin				
Asian	12 (6%)	3 (3%)	4 (12%)	12 (4%)
Black	36 (17%)	17 (19%)	8 (24%)	62 (19%)
Hispanic	1 (<1%)	0	0 (0%)	1 (<1%)
Indian subcontinent	139 (66%)	61 (69%)	11 (32%)	150 (46%)
Middle Eastern	3 (1%)	0	0 (0%)	11 (3%)
Mixed	2 (1%)	0	0 (0%)	7 (2%)
White	19 (9%)	8 (9%)	11 (32%)	83 (25%)
Unknown	0	0	0	0
Body-mass index, kg/m ²				
Median (IQR)	22.6 (20.2–25.4)	22.5 (20.8–25.2)	23.6 (20.5–30.5)	24.4 (21.2–27.9)
Range	15.7–48.5	14.6–42.2	12.7–45.2	14.8–47.2
Recent known tuberculosis contact	52 (25%)	17 (19%)	11 (32%)	70 (21%)
History of previous tuberculosis	8 (4%)	6 (7%)	0	41 (13%)
Comorbidities				
HIV positive	9 (4%)	11 (12%)	0	25 (8%)
None	147 (69%)	59 (66%)	14 (41%)	141 (43%)
Asthma	12 (6%)	4 (4%)	4 (12%)	38 (12%)
Diabetes	12 (6%)	3 (3%)	5 (15%)	39 (12%)
Other	41 (19%)	24 (27%)	17 (50%)	138 (42%)
Symptoms				
Cough	142 (67%)	47 (53%)	16 (47%)	238 (73%)
Fever	106 (50%)	40 (45%)	6 (18%)	144 (44%)
Night sweat	111 (52%)	47 (53%)	15 (44%)	152 (46%)
Weight loss	128 (60%)	46 (52%)	12 (35%)	154 (47%)
Haemoptysis	23 (11%)	8 (9%)	1 (3%)	53 (16%)
Lethargy	113 (53%)	49 (55%)	17 (50%)	168 (51%)
Other symptoms	131 (62%)	54 (61%)	21 (62%)	144 (44%)

Data are n (%), unless otherwise specified.

Table 1: Demographic and clinical characteristics of analysed participants

suggestive of tuberculosis unlikely to be caused by another disease, no positive microbiological culture available, the attending clinician decision to treat, appropriate response to therapy, and supportive histological evidence if available. Categories 1 and 2 combined are referred to as all active tuberculosis. In category 3, patients had a clinically indeterminate diagnosis and tuberculosis was neither highly probable nor reliably excluded.¹⁷ Patients assigned to category 4, other diseases, had no evidence of active tuberculosis using a combination of symptoms, risk factors, and negative bacteriology and evidence of another diagnosis if available.

Statistical analysis

The primary objective of this study was to derive transcriptomic signatures from a cohort of patients presenting with clinically suspected tuberculosis in routine clinical care and assess their performance at differentiating tuberculosis from its naturally occurring differential diagnoses. Category 3 participants were excluded from analyses, because they could not be reliably assigned tuberculosis or other diseases status. All others were included in the primary analysis. A summary of the methods used to derive novel signatures from the cohort is presented in the appendix (pp 1–3, 6–7).

80% of the cohort was randomly selected as a derivation set using the sample function in R (version 4.0.1). Differentially expressed transcripts between active tuberculosis and other diseases groups in the derivation set were used to find the signatures with the lowest classification errors using Random Forest, Leaps (version 3.1), and linear discriminant analysis (MASS R package version 7.3-53). These signatures were then trained in the same derivation set and tested in the remaining validation set containing 20% of the cohort. To estimate the prediction uncertainty of the model and report it as error bars on the receiver operating characteristic curves, the training and testing of the new signatures were repeated 1000 times using Monte Carlo cross-validation (resampling without replacement by use of the bootstrap and sample R function), which randomly created 1000 new 80:20 splits of the full dataset. A similar strategy has been used previously.^{6,7}

To address the possibility that transcriptomic signatures of host response in paucibacillary culture-negative highly probable tuberculosis might substantially differ from those in culture-positive tuberculosis, we derived three additional signatures (4, 5, and 6) by applying the same method specifically to highly probable tuberculosis and a randomly selected equal number of cases of other diseases.

The secondary objective of this study was to assess the performance of previously published transcriptomic signatures in this cohort. Three criteria suggestive of high diagnostic performance were used to identify the most rigorously derived previously published transcriptomic signatures to apply to our cohort: published risk score calculable in our cohort, comparisons of tuberculosis against other diseases (not healthy controls), and derivations from a sufficient dataset (more than 400 participants). Existing transcriptomic signatures published as transcript lists without coefficients (thus requiring retraining in their own dataset before application) were not useable. Conversely, training of previously published transcript lists in our cohort would create a new signature, not representing the original. Furthermore, retraining of a signature before use in a new population is not feasible in clinical practice. Three previously published transcriptomic signatures met all three criteria^{6–8} and were applied to 100% of the same

whole cohort and subgroups as the novel signatures to derive an area under the receiver-operating characteristic curves (AUC-ROC), sensitivity, and specificity.

Receiver operating characteristic analysis (ROCR package version 1.0-11 and EasyROC web-tool) was used to assess the diagnostic performance of the signatures in classifying patients against the final assigned diagnoses: category 1, category 2, and category 4. Where cases were significantly outnumbered by other diseases controls, a randomly selected size-matched sample of other diseases controls was used in the analysis. Sensitivity and specificity were calculated with the Youden Index to equally prioritise the two, then high sensitivity or specificity were selected to match with WHO targets for a novel confirmatory test.² Although the WHO criteria were designed to guide novel diagnostics for medium-to-high-incidence settings, they can also provide a relevant benchmark for performance of new tests in low-incidence settings. To account for the small possibility that a proportion of the group with highly probable tuberculosis had been misdiagnosed by our classification system, we did a sensitivity analysis as described in Kaforou et al,⁶ calculating diagnostic accuracies when a true prevalence of tuberculosis of 80%, 85%, or 90% was assumed.

The IDEA cohort size was powered to detect a 10% difference in diagnostic sensitivity between the two commercially available IGRAs. Accounting for the paired nature of the data and assuming independence of errors, 855 patients (after loss to follow-up, withdrawal, or exclusion because of missing or invalid index or reference test results) were required to detect this difference at the 5% significance level (two-tailed) with 90% power, based on a predicted 40% prevalence of active tuberculosis in the study population. From the resulting overall IDEA cohort, we used every available RNA sample from all patients who completed follow-up and who consented to genomic analysis.¹⁶

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study, and the corresponding author had final responsibility for the decision to submit for publication.

Results

Between Nov 25, 2011, and Dec 31, 2013, 1162 patients with suspected tuberculosis were enrolled, and 628 (54%) were included in the final analysis (figure 1). Of these 628, 301 (48%) had active tuberculosis, 212 (34%) had culture-confirmed tuberculosis, and 89 (14%) had highly probable tuberculosis (76 [85%] with culture-negative tuberculosis and 13 [15%] with cultures not done). The remaining 327 (52%) patients had other diseases (figure 1). As planned, the small subgroup of patients with clinically indeterminate disease was excluded from analysis, because tuberculosis as a cause for their symptoms could neither

	All tuberculosis (n=301)	Other diseases* (n=327)
Culture-negative tuberculosis	76 (25%)	..
Smear-positive tuberculosis	50 (17%)	..
Smear-negative tuberculosis	198 (66%)	..
Smear not tested	53 (18%)	..
Pulmonary tuberculosis†	101 (34%)	..
Extrapulmonary tuberculosis	164 (54%)	..
Pulmonary tuberculosis plus extrapulmonary tuberculosis	36 (12%)	..
Site of tuberculosis infection‡		
Abdomen	8 (3%)	..
Bone	6 (2%)	..
CNS	5 (2%)	..
Chest	2 (<1%)	..
Lung	136 (45%)	..
Lymph node	131 (44%)	..
Miliary tuberculosis (disseminated)	11 (4%)	..
Pericardium	5 (2%)	..
Pleura	23 (8%)	..
Spine	16 (5%)	..
Other site	22 (7%)	..
Pneumonia	..	69 (21%)
Sarcoidosis	..	34 (10%)
Cancer	..	26 (8%)
Non-pneumonia lower respiratory tract infection	..	24 (7%)
Bronchiectasis	..	21 (6%)
Reactive lymphadenopathy	..	14 (4%)
Upper respiratory tract infection	..	10 (3%)
Asthma	..	10 (3%)
Atypical <i>Mycobacterium</i> spp infection	..	9 (3%)
Other diagnoses§	..	110 (34%)

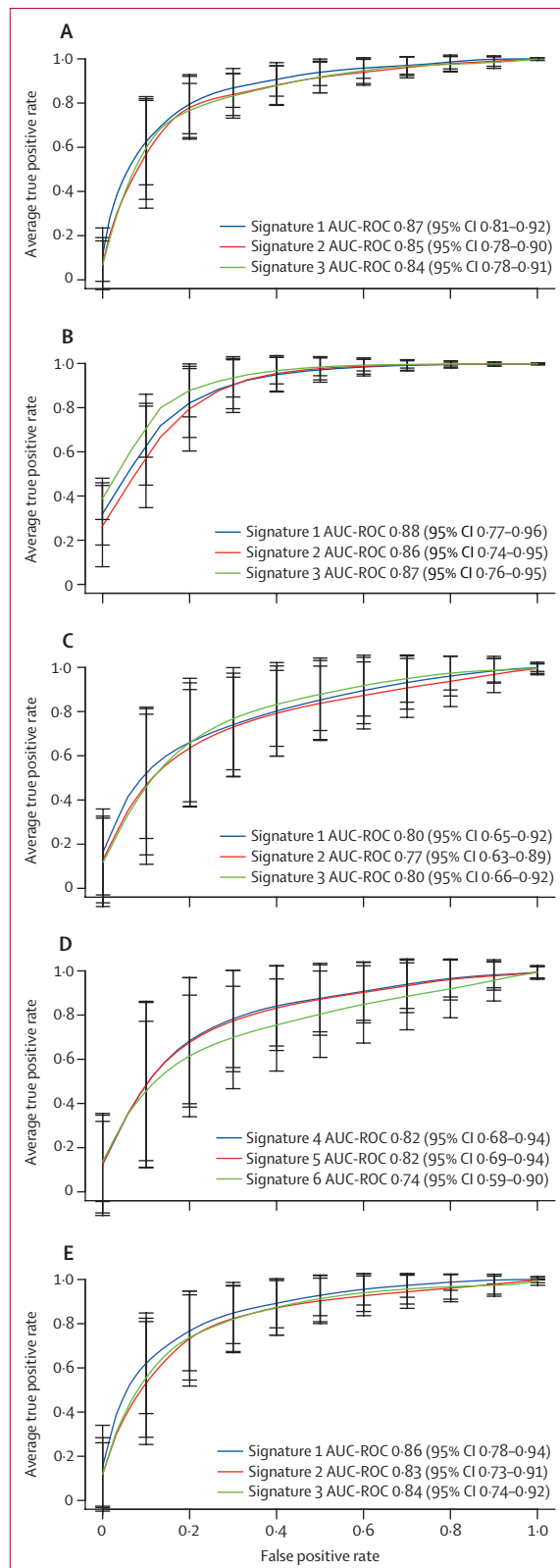
*Some patients had multiple diagnoses. †For patients with highly probable tuberculosis, 16 (18%) had pulmonary tuberculosis, 64 (72%) had extrapulmonary tuberculosis, and nine (10%) had both. ‡Some patients had tuberculosis at multiple anatomical sites. §Fewer than five cases per diagnosis.

Table 2: Final presenting diagnoses of patients with tuberculosis and other diseases

be ruled in nor ruled out (n=34). Demographic and clinical characteristics of the cohort are summarised in table 1.

Of the 301 patients with active tuberculosis, 50 (17%) were sputum smear-positive, 101 (34%) had pulmonary tuberculosis, 164 (54%) had extrapulmonary tuberculosis, and 27 (13%) had both (table 2). The proportions of these diagnoses were comparable with the proportions of pulmonary tuberculosis and extrapulmonary tuberculosis in England as a whole.¹⁹ Patients with other diseases comprised a diverse group (table 2). 12 (2%) of the 628 patients included in the final analysis were taking immunosuppressant medication at enrolment.

Transcriptomic analysis of 47 275 transcripts in whole blood showed 171 differentially expressed transcripts between patients with active tuberculosis and those with other diseases (appendix pp 9–13). The new signatures



were derived in a randomly selected 80% and tested in a distinct 20% of the full ($n=628$) cohort for all patients with active tuberculosis (regardless of tuberculosis subgroup). The performance of these signatures was then tested in the different subgroups of tuberculosis (figure 2; table 3). The best-performing diagnostic signature derived from the whole cohort was signature 1, comprising 13 transcripts, with an AUC-ROC of 0.87 (95% CI 0.81–0.92), sensitivity of 77% (66–87), and specificity of 84% (74–91) for identification of active tuberculosis. In patients with culture-confirmed pulmonary tuberculosis, it had an AUC-ROC of 0.88 (0.77–0.96), sensitivity of 82% (65–95), and specificity of 83% (68–92), and in those with highly probable tuberculosis, these values were 0.86 (0.71–0.95), 77% (56–94), and 77% (57–95; figure 2; table 3).

Diagnostic performance of signatures 1–3 was also assessed in patients with extrapulmonary tuberculosis (figure 2; table 3). Signature 1 performed best, with an AUC-ROC of 0.86 (95% CI 0.78–0.94), sensitivity of 76% (61–91), and specificity of 82% (68–94).

For signatures derived exclusively in the highly probable tuberculosis subgroup, signature 4, comprising one transcript (*GBP5*), performed best and provided an AUC-ROC of 0.82 (95% CI 0.68–0.94), sensitivity of 65% (42–87), and specificity of 84% (68–100). Signature 5 had an AUC-ROC of 0.82 (0.68–0.94) and signature 6 an AUC-ROC of 0.74 (0.59–0.90; appendix p 14).

The best-performing previously published signature for detecting active tuberculosis, culture-positive pulmonary tuberculosis, highly probable tuberculosis, and extrapulmonary tuberculosis was derived by Sweeney and colleagues.⁸ For detection of active tuberculosis, it had an AUC-ROC of 0.83 (95% CI 0.80–0.86), sensitivity of 78% (73–83), and specificity of 76% (70–80; table 3). For detecting highly probable tuberculosis, it had results of 0.71 (0.64–0.79), 76% (66–85), and 62% (51–72; table 3).

None of the relevant cohort-derived signatures or previously published signatures achieved minimal targets of paired sensitivity (98% for smear-positive culture-positive pulmonary tuberculosis and 65% for smear-negative culture-positive pulmonary and extrapulmonary tuberculosis) and specificity (98%) set by the WHO target product profiles for a non-sputum-based diagnostic test. These targets and the performance of the

Figure 2: Performance of the cohort-derived signatures in the whole cohort and tuberculosis subgroups

(A) Active tuberculosis ($n=301$) versus other diseases ($n=327$). (B) Culture-confirmed pulmonary tuberculosis ($n=112$) versus other diseases ($n=112$). (C) Highly probable tuberculosis ($n=89$) versus other diseases ($n=89$) using signatures 1–3, which were derived from the full all active tuberculosis cohort. (D) Highly probable tuberculosis ($n=89$) versus other diseases ($n=89$) using signatures 4–6, derived from the highly probable tuberculosis subgroup. (E) Extrapulmonary tuberculosis ($n=164$) versus other diseases ($n=164$). Error bars represent the 95% CI. AUC-ROC=area under the receiver-operating characteristic curve.

Gene signature	Comparison groups	Previously reported performance of published signatures				All active tuberculosis vs other disease (n=628)				Culture-confirmed pulmonary tuberculosis vs other disease (n=224)				Highly-probable tuberculosis vs other disease (n=178)				Extrapulmonary tuberculosis vs other disease (n=328)			
		Cohort site (size) ^{††}	AUC-ROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUC-ROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUC-ROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUC-ROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUC-ROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)				
Signature 1	13 transcripts ^{‡§}	0.87 (0.81–0.92)	0.77 (0.66–0.87)	0.84 (0.74–0.91)	0.88 (0.77–0.96)	0.82 (0.65–0.95)	0.83 (0.68–0.96)	0.75 (0.52–0.93)	0.80 (0.66–0.93)	0.72 (0.47–0.92)	0.75 (0.52–0.93)	0.86 (0.78–0.94)	0.76 (0.61–0.91)	0.82 (0.68–0.94)			
Signature 2	20 transcripts ^{‡§}	0.85 (0.78–0.90)	0.77 (0.65–0.87)	0.82 (0.72–0.90)	0.86 (0.74–0.95)	0.75 (0.63–0.90)	0.82 (0.63–0.96)	0.74 (0.53–0.93)	0.77 (0.63–0.89)	0.68 (0.44–0.88)	0.74 (0.53–0.93)	0.83 (0.73–0.91)	0.78 (0.61–0.92)	0.78 (0.62–0.93)			
Signature 3	FCGR1C, GBP5, UB2L6	0.84 (0.78–0.91)	0.75 (0.64–0.85)	0.81 (0.72–0.90)	0.87 (0.76–0.95)	0.80 (0.62–0.95)	0.78 (0.62–0.95)	0.79 (0.56–0.94)	0.80 (0.66–0.92)	0.70 (0.47–0.89)	0.79 (0.56–0.94)	0.84 (0.74–0.92)	0.74 (0.58–0.88)	0.79 (0.66–0.93)			
Kaforou et al ⁶	Disease Risk Score 44 transcripts vs other diseases	Malawi (n=471)	1 (1–1)	1 (1–1)	0.96 (0.93–1)	0.81 (0.78–0.85)	0.74 (0.69–0.79) [†]	0.77 (0.72–0.81)	0.85 (0.80–0.90)	0.75 (0.66–0.83) [†]	0.87 (0.79–0.92)	0.75 (0.61–0.85) [†]	0.70 (0.63–0.78)	0.76 (0.66–0.85) [†]	0.61 (0.50–0.71)	0.79 (0.74–0.84)	0.73 (0.66–0.80) [§]	0.78 (0.71–0.84)			
Anderson et al ⁶	Disease Risk Score 51 transcripts	South Africa, Malawi, and Kenya (n=503)	0.89 (0.82–0.95)	0.83 (0.69–0.94)	0.83 (0.73–0.92)	0.68 (0.64–0.72)	0.75 (0.70–0.80) ^{¶¶}	0.50 (0.46–0.58)	0.67 (0.60–0.74)	0.74 (0.65–0.82) ^{¶¶}	0.51 (0.41–0.60)	0.56 (0.45–0.67)	0.63 (0.55–0.71)	0.60 (0.49–0.70) ^{¶¶}	0.56 (0.45–0.67)	0.69 (0.63–0.75)	0.77 (0.70–0.83)	0.52 (0.44–0.60)			
Sweeney et al ⁸	TB risk score (culture-positive or smear-positive) vs other diseases ^{††}	0.83 (0.80–0.86)	0.78 (0.73–0.83) ^{##}	0.76 (0.70–0.80)	0.86 (0.81–0.91)	0.75 (0.66–0.83) ^{##}	0.88 (0.81–0.94)	0.62 (0.51–0.72)	0.71 (0.64–0.79)	0.76 (0.66–0.85) ^{##}	0.62 (0.51–0.72)	0.83 (0.78–0.87)	0.85 (0.79–0.90) ^{§§}	0.70 (0.63–0.77)			
German validation cohort	..	Germany (n=1067)	0.75 (0.65–0.86)	0.50	0.61			
Kenyan validation cohort	..	Kenya (n=1180)	0.91 (0.88–0.95)	0.77	0.86			
South African validation cohort	..	Malawi and South Africa (n=1357)	0.82 (0.79–0.84)	0.69	0.74			

AUC-ROC=area under the receiver-operating characteristic curve. *Reported size of cohorts includes the derivation cohort (comprising a training set and test set) and the validation cohort. [†]Youden Index cutoff was 110.21, 106.00, and 108.60. [‡]See appendix (pp 4–5) for full transcript lists. [§]Youden Index cutoff was 108.36. [¶]Youden Index cutoff was 92.18, 92.38, and 91.44. ^{||}Youden Index cutoff was 91.44. ^{**}The Sweeney et al⁸ signature was derived using the combination of three discovery cohorts comprising 1023 participants, then applied separately to three validation cohorts (each reported in table with size including the 1023 shared discovery cohorts). ^{††}Other comparisons not reported in this table include active tuberculosis versus healthy controls and active tuberculosis versus latent tuberculosis infection. ^{##}Youden Index cutoff was –2.82, –3.48, and –3.23. ^{§§}Youden Index cutoff was –3.48.

Table 3: Diagnostic performance of cohort-derived signatures and previously published signatures in patients with tuberculosis diagnosis

Table 3: Diagnostic performance of cohort-derived signatures and previously published signatures in patients with tuberculosis diagnosis

novel signatures against them are reported in the appendix (pp 8, 15).

When the highly probable tuberculosis group was assigned true prevalence of 90%, the sensitivity of signature 1 increased to 77% (95% CI 47–100), at a prevalence of 85% it was 80% (47–100), and at a prevalence of 80% it was 84% (47–100; appendix p 14).

Discussion

In this large, multicentre, prospective cohort study in routine practice in a low-incidence setting that covered the full clinical spectrum of active tuberculosis and its differential diagnoses with concomitant latent tuberculosis infection, transcriptomic signatures showed suboptimal diagnostic accuracy, bringing their clinical utility into question. The best-performing cohort-derived signature for all tuberculosis cases and the best previously published signature for all tuberculosis categories⁸ had insufficient diagnostic accuracies, sensitivities, and specificities.

Signatures derived from our clinically relevant real-life cohort by multiple methods and in key clinical tuberculosis subgroups performed only marginally better than pre-existing signatures, with none achieving an AUC-ROC of 0·90 in any clinical category, suggesting that a ceiling has been reached in the diagnostic potential of transcriptomic signatures in this setting. No signatures improved on the performance of IGRAs, with T-SPOT.TB achieving a sensitivity of 81% and specificity of 85% for active tuberculosis in this cohort.¹⁶ Given that IGRAs are not recommended for the diagnosis of active tuberculosis because of insufficient diagnostic accuracy,^{16,20} it follows that transcriptomic signatures also cannot be recommended.

Two of the most promising previously published signatures were derived in case-control studies comparing culture-positive tuberculosis and other diseases.^{7,8} This could explain their reduced accuracy in our cohort and their poorer performance in identifying highly probable tuberculosis relative to culture-positive pulmonary tuberculosis. Another previously published signature was derived and validated by Anderson and colleagues⁶ in a large group of children with suspected tuberculosis (discovery cohort $n=346$, validation cohort $n=157$) in high-incidence, low-resource settings. This signature achieved an AUC-ROC of 0·89 (95% CI 0·82–0·95), sensitivity of 83% (67–94), and specificity of 83% (73–92). Its reduced performance in our cohort probably reflects differences in the study populations; children might have a distinct transcriptomic response to tuberculosis. Additionally, the differential diagnoses of non-tuberculosis cases in their child cohort differ greatly from adults; there were no cases of cancer or sarcoidosis in these children, whereas malnutrition, malaria, and helminth infections were frequent. To address the question of whether other signatures might perform better, the transcriptomic data from our cohort will be publicly available to enable future investigators to assess

the performance of other published or novel signatures in our cohort. Culture-unconfirmed, highly probable tuberculosis and extrapulmonary tuberculosis have unmet diagnostic needs, for which hope has been pinned on non-sputum-based host-response biomarkers, particularly transcriptomics. Our study provides the first assessment of this technology in these key subgroups of adults with tuberculosis. The two previously published signatures derived exclusively from culture-confirmed pulmonary tuberculosis performed markedly less well in diagnosing highly probable tuberculosis. The novel signatures showed inferior performance for identification of highly probable tuberculosis than for culture-confirmed tuberculosis, even when we derived signatures specifically from and for that subgroup.

A strength of our study is the rigorous case definition for highly probable tuberculosis that used a validated composite reference standard, including appropriate response to therapy after 6 months.^{16,17} The need to define a composite reference standard demonstrates the diagnostic challenge of culture-negative tuberculosis in this setting and hence the need for a rapid, accurate test. Moreover, clinically indeterminate patients, in whom tuberculosis could not be ruled out but who did not meet the composite reference standard for highly probable tuberculosis, comprised only a small fraction (5%) of the cohort and were excluded from analysis as planned. However, a proportion of patients with highly probable tuberculosis could conceivably have been misclassified, which might falsely lower the accuracy of the signatures. We tested this hypothesis with a sensitivity analysis, in which the prevalence of tuberculosis cases in the highly probable tuberculosis group was set to 80%, 85%, and 90%. The results showed that even when 20% of the patients were assumed to be misclassified by our composite reference standard, improvement in the signature's performance was inadequate.

In our cohort, no novel or previously published signatures achieved the minimal WHO sensitivity (65%) and specificity (98%) cutoffs for a non-sputum-based confirmatory test for sputum smear-negative tuberculosis, using culture-confirmed tuberculosis as a gold standard. This finding aligns with studies^{21,22} of previously published signatures in cohorts of 181 and 293 patients with suspected pulmonary tuberculosis in very-high-incidence settings, which showed that the signatures did not meet these WHO requirements.

Thus, we found that in a routine practice setting where tuberculosis must be differentiated from other diseases with similar clinical presentations, transcriptomic signatures did not show sufficient diagnostic accuracy to be clinically useful. This finding might be because a positive transcriptomic signature is a marker of inflammation in response to a foreign stimulus, rather than a marker of the specific pathogen, making its diagnostic specificity inadequate to deliver a positive result reliable

enough to initiate tuberculosis therapy. Diagnosis of an infectious disease using antigen or immune response to antigen, however, relies on pathogen specificity. Future work on transcriptomic signatures should perhaps therefore combine them with antigen-stimulated responses.

This study had some limitations. Our findings cannot be extrapolated to children or to high-incidence, low-resource settings, although two smaller studies^{21,22} of patients with suspected pulmonary tuberculosis in South Africa have also suggested insufficient diagnostic accuracy. Additionally, owing to the low prevalence of HIV-associated tuberculosis in our cohort (20 [7%] of 301 patients), conclusions cannot be drawn about the use of transcriptomics in this co-infected group, although there is no reason to believe performance would be better in the presence of HIV co-infection. To reduce the disproportionately large number of HIV-positive patients with other diseases, a subgroup of these patients were selected randomly. A similar method was used by Anderson and colleagues⁶ and would not be expected to alter the validity of the overall results. Although RNA sequencing is superseding the use of microarrays, a microarray was used here to enable parallel assessment of previously published signatures, all of which were developed with microarrays. Because RNA sequencing also quantifies abundance of transcripts, unstimulated whole-blood RNA sequencing would probably not perform substantially better. The small proportion of participants taking immunosuppressant medication, 2% of the final cohort, is unlikely to have affected our analysis. The distribution of ethnicity in our cohort is representative of patients with suspected tuberculosis in routine clinical practice in England.¹⁹

In summary, these strong, generalisable data do not support the use of transcriptomic signatures as diagnostic tools for tuberculosis in routine practice in low-incidence, high-resource settings.

Contributors

AL conceived the study and designed it with YT, JJD, and OMK. PB and AL supervised the study. LTH did the microarray experiments and analysed microarray data. MT-W collected data, MT-W and AH managed data, and TDP, MT-W, and AH analysed data. LTH and TDP accessed and verified data and PB had oversight of transcriptome data. LTH, PJ, UN, YT, and JJD did the statistical analysis. LTH, TDP, LCB-A, OMK, and AL wrote the manuscript, which was reviewed by YT, JJD, and PB.

Declaration of interests

AL reports issued patents underpinning IGRA and next-generation IGRA, some of which were assigned by the University of Oxford to Oxford Immunotec, resulting in royalty entitlements for the University of Oxford and AL. AL is a named inventor on the following patents EP05729257.5, EP1735623[B1], US8,105,797[B2], EP2069792, EP2069792[B1], EP2005182, EP2005182[B1], US8,765,336[B2], EP10716313.1, EP2417456[B1], US9,377,460[B2], US9360480[B2], EP0941478[B2], EP1152012[B1], EP1735623[B1], US8105797[B2], EP1144447[B1], and US9005902[B2]. All other authors declare no competing interests.

Data sharing

All the data used for this report, including raw and normalised microarray data and patients' clinical data, will be made available through the GEO repository (accession number GSE144127).

Acknowledgments

The study was funded by the National Institute for Health Research (NIHR) Efficacy, Mechanisms, and Evaluation Programme, grant number 12/65/27 (Validation of New Technologies for Diagnostic Evaluation of Tuberculosis [VANTDET]). AL and JJD are UK NIHR Senior Investigators Emeriti. YT and JJD are supported by the NIHR Birmingham Biomedical Research Centre. TDP was supported by the Research England Global Challenge Research Fund. The Imperial College NIHR Health Protection Research Unit in Respiratory Infections and the Imperial College Biomedical Research Centre are also acknowledged for their support. We are grateful to the VANTDET independent scientific advisory group, comprising Madhad Noursedeghi (University College London, London, UK), David Connell (University of Dundee, Dundee, UK), Sanjeev Krishna (St George's University of London, London, UK), and Kevin Fennelly (National Heart, Lung, and Blood Institute, Bethesda, MA, USA) for critical appraisal of the manuscript. We thank all the patients for their participation in the study and the medical staff and nursing at the recruiting hospitals. We would like to thank Philippe Froguel (Institut de Biologie de Lille, Lille, France) whose laboratory did the microarray analyses. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

References

- 1 WHO. Global tuberculosis report 2019. World Health Organization, 2019. <https://www.who.int/publications/i/item/global-tuberculosis-report-2019> (accessed Nov 25, 2020).
- 2 WHO. High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting. World Health Organization, 2014. <https://www.who.int/tb/publications/tpp-report/en/> (accessed Nov 25, 2020).
- 3 Pai M, Schito M. Tuberculosis diagnostics in 2015: landscape, priorities, needs, and prospects. *J Infect Dis* 2015; **211** (suppl 2): S21–28.
- 4 Kohli M, Schiller I, Dendukuri N, et al. Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. *Cochrane Database Syst Rev* 2018; **8**: CD012768.
- 5 Unitaid. Tuberculosis diagnostics technology landscape. Unitaid, 2017. <https://unitaid.org/assets/2017-Unitaid-TB-Diagnostics-Technology-Landscape.pdf> (accessed Nov 25, 2020).
- 6 Anderson ST, Kaforou M, Brent AJ, et al. Diagnosis of childhood tuberculosis and host RNA expression in Africa. *N Engl J Med* 2014; **370**: 1712–23.
- 7 Kaforou M, Wright VJ, Oni T, et al. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. *PLoS Med* 2013; **10**: e1001538.
- 8 Sweeney TE, Braviak L, Tato CM, Khatri P. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med* 2016; **4**: 213–24.
- 9 Berry MPR, Graham CM, McNab FW, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 2010; **466**: 973–77.
- 10 Maertzdorf J, McEwen G, Weiner J 3rd, et al. Concise gene signature for point-of-care classification of tuberculosis. *EMBO Mol Med* 2016; **8**: 86–95.
- 11 Roe JK, Thomas N, Gil E, et al. Blood transcriptomic diagnosis of pulmonary and extrapulmonary tuberculosis. *JCI Insight* 2016; **1**: e87238.
- 12 Warsinske H, Vashisht R, Khatri P. Host-response-based gene signatures for tuberculosis diagnosis: a systematic comparison of 16 signatures. *PLoS Med* 2019; **16**: e1002786.
- 13 Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; **282**: 1061–66.
- 14 Rutjes AWS, Reitsma JB, Di Nisio M, Smidt M, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ* 2006; **174**: 469–76.
- 15 Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004; **140**: 189–202.
- 16 Whitworth HS, Badhan A, Boakye AA, et al. Clinical utility of existing and second-generation interferon- γ release assays for diagnostic evaluation of tuberculosis: an observational cohort study. *Lancet Infect Dis* 2019; **19**: 193–202.

- 17 Dosanjh DPS, Hinks TS, Innes JA, et al. Improved diagnostic evaluation of suspected tuberculosis. *Ann Intern Med* 2008; **148**: 325–36.
- 18 Hoang LT, Lynn DJ, Henn M, et al. The early whole-blood transcriptional signature of dengue virus and features associated with progression to dengue shock syndrome in Vietnamese children and young adults. *J Virol* 2010; **84**: 12982–94.
- 19 Public Health England. Tuberculosis in England: 2018. UK Government, 2019. <https://www.gov.uk/government/publications/tuberculosis-in-england-annual-report> (accessed Nov 25, 2020).
- 20 Internal Clinical Guidelines Team. Tuberculosis: prevention, diagnosis, management and service organisation. London: National Institute for Health and Care Excellence, 2016.
- 21 Turner CT, Gupta RK, Tsaliki E, et al. Blood transcriptional biomarkers for active pulmonary tuberculosis in a high-burden setting: a prospective, observational, diagnostic accuracy study. *Lancet Respir Med* 2020; **8**: 407–19.
- 22 Penn-Nicholson A, Mbandi SK, Thompson E, et al. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci Rep* 2020; **10**: 8629.